Shravya Kanchi

☎ (540) 449-7120 | ✉ shravya@vt.edu | shrave.github.io | github.com/shrave | Google Scholar

Education

Virginia Polytechnic Institute and State University (Virginia Tech) Blacksburg, VA
Ph.D. in Computer Science, GPA 3.76/4.0 Advisor: Dr. Bimal Viswanath *Aug. 2021 – Present*

IIIT Hyderabad India
MS by Research in Computer Science and Engineering, GPA 9/10.0 *Jul. 2018 – Aug. 2021*

IIIT Sricity India
B.Tech in Computer Science and Engineering with Honors, GPA 8.79/10.0 *Jul. 2014 – May. 2018*

Selected Projects

Synthetic data to strengthen security defenses Under Submission

- Conducted a comprehensive measurement study, identifying three major data challenges impacting security defenses by analyzing 35 top-tier security research articles.
- Engineered a VQVAE-based generative AI model for synthetic tabular data creation, driving a 20% performance improvement across five distinct security defenses.
- Showed superior performance gains as compared to 5 SOTA tabular data generators on security classifiers.

Toxicity mitigation in chatbot conversational training pipelines Under Submission

- Proposed a framework for toxicity mitigation on diverse chatbot customization pipelines.
- Proposed context-aware and context-agnostic approaches leveraging LLMs to identify toxic conversations which outperformed industry API services in a data poisoning setting.
- Safety-aligned chatbots by adding synthetically crafted conversations with desirable traits through Direct Preference Optimization (DPO) achieving near-zero toxicity.

Multi-perspective Access Control System CODASPY'21

- Designed a smart home privilege system for multi-user, multi-device environments, ensuring least privilege access.
- Implemented a formal access control specification in a Python Flask server.

Publications

[**Submitted to USENIX Security'25**] "Title anonymized for double-blind submission." 1st **author**.
[**Submitted to USENIX Security'25**] "Title anonymized for double-blind submission." 1st **author**.
[**IEEE S&P'24**] "An Analysis of Recent Advances in Deepfake Image Detection in an Evolving Threat Landscape." 3rd author.
[**ACSAC'23**] "First Look at Toxicity Injection Attacks on Open-domain Chatbots." 3rd author.
[**WWW Workshop'22**] "SEBI Regulation Biography" 3rd author.
[**CODASPY'21**] "A multi perspective access control in a smart home." 1st **author**.

Work Experience

Graduate Research Assistant Jan. 2021 – Jun. 2021
Collaboration between IIIT Hyderabad & JP Morgan Chase

- Built the first named-entity labeled corpus for SEBI regulations, with 7,500 sub-regulations.
- Proposed 7 unique named entity types specific to the Indian securities regulatory context.
- Developed an Overlapping Named Entity Recognition tool for SEBI with 87.47% precision.

Technical Skills

GenAI: LLMs, LoRA, model customization, DPO, Stable Diffusion, VAE, GAN, prompt engineering, hyper-parameter search, BERT, Fine-tuning, PEFT, AutoML.

Security: ML-based malware detection, Phishing, Spam, Network IDS, concept drift, BGP, website privacy.

Frameworks: Deep learing - Pytorch, Tensorflow, Keras, Flask, MongoDB

Developer Tools: Git, Linux, VS Code, Vim, Jupyter Notebook, Docker, Nmap, label-studio

Libraries: NumPy, Huggingface, Transformers, RayTune, Matplotlib, gnuplot, Scikit-Learn, pandas, LightGBM, SpaCy

Languages: Python, C++, C, Shell, SQL